

Extension of Victimization in Unsolicited E-mail Messages with Attachments (UEMAs): An Explanation of Seeking and Exposing Process

ABSTRACT

While the growing scale of Internet use brings about great convenient for users, phenomena of unsolicited e-mail pose new threats and challenges. Previous literature was concentrated on general analysis of such messages, leaving many particular respects untouched. This study focuses on the extension of victimization of unsolicited messages e-mail with attachments (UEMAs). Based on the analysis of two samples, one comprised of 501 (sampling done in May 2006), and the other comprised of 490 (sampling done in March 2008),pieces of UEMAs, the study finds that e-mail account exposing and seeking can both contribute to victimization; while receiving of unsolicited messages is the initial victimization, reading and reacting to messages could lead to additional victimization from virus attack or financial fraud, and from conspiracy in illegitimate operations such as tax evasion or transaction of falsified documents.

Keywords: E-Commerce, Survey, Cybercrime, Unsolicited E-mail Messages with Attachments (UEMAs), Victimization, Conspiracy

Hacking all the way into spam

The twentieth century witnessed many outstanding creations of human beings, of which the Internet is one extending its power to today. Radical group (2005) found that more than 683 million users held 1.2 billion e-mail accounts in communication and marketing. It implied that, besides other things, each user could have more than one account. In practice, many frequent users have two or more accounts. Despite the convenience that this unprecedented tool creates, unsolicited e-mail message, its by-product, becomes a nuisance that each user meets by chance. Recipients' property rights, fair trade, public morals, cybersecurity, data protection, as well as other content-related and goods-related transgresses and offences all pose great challenges and threats to society that is symbolized by cyberspace (Li 2006).

Many people have no interest in talking about this issue. Even some economists, whom I had conversation with in a conference in 2005, ignored the significance of such a research. Their straightforward logic was that “unsolicited e-mail is everyday phenomena; everyone knows something about it; and no one thinks it so serious.” One of the economists told me that he received spams everyday and deleted them in a few seconds without extra trouble. They tended to devaluate the academic subject by ignoring the impact of the research object through raising examples in their own life, which seemed to be strong enough proof and sound enough understanding.

However, numerous authors have scrutinized the phenomenology of unsolicited commercial e-mail (UCE, or unsolicited business e-mail, UBE, or simply spam) from points of view of economics, commerce, law, technology, sociology, culture, and so on. Others have principally dealt with costs and benefits of senders and recipients derived from unsolicited e-mails (Khong 2001, and 2004), general impact on productivity of individual employees and enterprises (Nucleus 2003, 2004), scale and volume of unsolicited e-mails (Radical Group 2005), impact on consumers’ attitudes and confidence towards e-commerce (TACD 2003; Fallows 2003; Harris Interactive 2003), higher possibility of receiving unsolicited e-mail due to online-published addresses (Federal Trade Commission 2002b), ignorance of removal requests by senders (Federal Trade Commission 2002a), and technical and legal solution on unsolicited e-mail (Sorkin 2001).

Few studies, nevertheless, have touched on unsolicited e-mail messages that have attachments, particularly on what kind of risks a single e-mail user might face. Li (2007) was the first known study specifically on such a phenomenon, based on a sample of 501 attachments (sampling done in May 2006). In this study, two samples were used, one is the same sample used in Li (2007), and the other is comprised of 490 pieces (sampling done in March 2008), of unsolicited e-mail messages with attachment, presenting the first analysis of types and validity of sender column, types and validity of subject column, types of offers of message content, and types, sizes and nature of attachments of these messages. In this paper, the term “spam” is deliberately taking into account the lack of a universally-accepted unified definition. At the same time, the author prefers the phrase of “unsolicited e-mail” without the

modifier “commercial” or “business”, in order to enlarge the coverage of UEMAs to virus spreaders in this study.

Literature review

At the same time as the augmented capacity of computers and networks to process information, “a wealth of information” could result in a “poverty of attention” (Simon 1982). Unsolicited business e-mail (UBE) or unsolicited commercial e-mail (UCE) incarnates an instance where e-mail users have to deal with redundant information they anticipate not to consume. Unsolicited e-mail gives rise to an unconstructive representation of e-marketing, alarming e-mail users from trusting e-mail communication.

Unsolicited e-mail sent out to multiple recipients has extensive unfavourable consequences on e-mail marketing. Karnell (2002) found that an increasing number of e-mails had never been read by recipients, and users prefer limiting or disrupting business contacts with these senders. The prevailingness of unsolicited messages exhibited an emergent intimidation to the information society (World Summit on the Information Society 2003, paragraph 37).

Spammer-X (2004) narrated his/her anecdote with reference to reasons and methods of spamming, revealing the astuteness of defeating anti-spam techniques, avoiding being identified, in addition to escaping the law. McWilliams (2005) accounted the world of spammers and spam-fighters, furnishing information on mechanisms of spamming and spam-fighting. Goodman (2004) presented the most typical spam traps and explained why existing solutions were ineffective. Lambert (2003) analysed a wide range of e-mails in attempt to generate a silhouette of spam and develop a profile of spammer.

The United States Federal Trade Commission (1998) speculated the issue from the standpoint of consumer protection, identifying a dozen of most likely spam scams, covering spam and scams from business opportunities, quick money, working at home, to guaranteed loans, and so on.

Li (2006) recapitulated six challenges that the spam brought to society: recipients' property rights, fair trade, public morals, cybersecurity, data protection, as well as other

content-related and goods-related transgresses and offences. From the standpoint of senders, all these challenges could be classified into two bigger categories: victim seeking and conspirator seeking. From the standpoint of recipients, they were confronted with preliminary victimization (being spammed), supplementary victimization (being defrauded, or attacked by viruses), or committing offences (tax evasion, or transaction and use of falsified documents).

Nucleus (2003) reported in-depth interviews with 117 employees and extensive interviews with 28 IT administrators. They found that an average of 13.3 unsolicited messages reached the employee per day; each employee has to waste an average of 6.5 minutes per day dealing with unsolicited messages. They calculated that unsolicited messages caused an average 1.4 percent of productivity loss per employee per year, tantamount to an average cost of 874 dollars per employee per year.

Nucleus (2004) reported further interviews with employees at 82 Fortune 500 companies. They found that users received an average of twice the number of previous year's unsolicited messages, with an average 3.1 percent of productivity loss in 2004. They also established that the function of technical solution to unsolicited messages became less efficacious.

Fallows (2003) reported the Pew Internet & American Life Project, which collected data from a national telephone survey of 2,200 adults and a compilation of more than 4,000 first-person narratives about unsolicited messages. Their findings showed that unsolicited messages caused some e-mail users to use e-mail less, and trusted the online environment less, while fear of unsolicited messages increased.

The deteriorated consumers' confidence on unsolicited e-mails results from the losing control over their own accounts. Federal Trade Commission (2002a) tested 215 addresses from spam with "remove me" claims, and found that unsubscribe demand was usually ignored. While senders of unsolicited e-mails do not provide effectual unsubscribe method, they are harvesting addresses from all over the Internet.

Federal Trade Commission (2002b) put 250 new, undercover e-mail addresses in 175 different locations on the Internet, including web pages, newsgroups, chat rooms, message boards, and online directions for web pages, instant message users,

domain names, resumes, and dating services. They found that web pages, newsgroups, and chat rooms were all attractive to unsolicited message senders. Federal Trade Commission (2005) found that spammers continued to harvest email address posted on web sites, and to a much lesser extent, those posted on blogs and USENET groups. Masking email addresses when posting on web sites could substantially reduce the risk of harvesting.

Federal Trade Commission (2003a) reported that 66 percent of spam messages were fraudulent in sender or subject columns, or in the message itself. False advertisements might distort the normal market of goods and services, harm the normal trade order, and reduce the consumers' confidence (The Directorate General of Telecommunications Ministry of Transportation and Communications 2005, pp. 6-7). Large volume of spams, malicious programs and malicious linkages contained in messages were main threats (PC World 2003).

Harris Interactive surveyed 2,376 adults online in 2003, and found that most online adults reported that they had received more spam than six months earlier. Only 14 percent have seen a decline in the volume of spam. A majority of respondents reported unsolicited messages annoying or very annoying (Taylor 2003).

Many spammers send messages by unauthorized use of accounts of other individuals or organizations (Organization of Economic Cooperation and Development 2004). E-mail addresses harvesting software can collect this information automatically from web pages (Boldt, Carlsson and Jacobsson 2004, p. 8). Based on discussion of spyware and findings from two experiments, Boldt, Carlsson and Jacobsson (2004, p. 4) concluded that spyware had a negative effect on computer security and user privacy. Spyware enables spreading of e-mail addresses that may result in receiving unsolicited e-mails.

Khong (2004) stated that although it was difficult to measure costs and benefits of the spammer, if the benefit obtained from the activity outweighs the cost, the spammer would carry out spamming activity. It follows that if there is one successful commercial transaction, the spammer can realize his or her benefit. The costs that are involved in the spamming can be roughly estimated according to costs of bandwidth, message sending, and obtaining of users' address. Costs of bandwidth and message sending are ignorable. According to Sadowsky

(2003, p. 55), the spammer could obtain users' e-mail address in 13 situations. Obviously, the most convenient and least expensive way is to harvest e-mail addresses automatically with specific software, which is also available from Internet, either being free of charge or with an inexpensive price.

The revenue of spammer from sending message has been found high in a few studies (Goodman and Routhwaite 2004). Cobb (2003, p. 2) suggested the concept of "the parasitic economics of spam," meaning that the act of sending a message cost the sender less than it cost all other parties impacted by sending of the message.

Costs and benefits of the spammed can also be estimated. Costs induced by spam to the spammed have a wide coverage, including waste of users' time, bandwidth and storage, cost of anti-spam solution, and cost of overloading at the mailbox (Gauthronet and Drouard 2001). The average time and money lost in processing a single message might not be so significant. However, theoretically, aggregate losses of time and money taken in dealing with these messages might be huge (Li 2006). Spam also induces costs of bandwidth and storage, losses in interruption of services, and anti-spam solutions (Gauthronet and Drouard 2001). Interruption of services is unfortunate for both providers and users in causing business, confidence, and other losses. Worldwide revenue for anti-spam solutions will exceed 1.7 billion dollars in 2008 (IDC 2005). If e-mail address has ever been put on institution's Web site, or personal homepage, it is highly possibly that the address will be harvested, sold, and abused by senders (Federal Trade Commission 2003b).


Finally, unsolicited e-mail is nothing useful and beneficial to recipients. From all of the previous studies, it is reasonable to conclude that recipients undertake pure losses, not only the monetary, but also the psychological (Li 2006).

Methodology

The paper presents a case study on UEMAs, analysing the sender column, subject column, content and attachments of these message. Two samples were used, one was composed of 501 messages, and the other was composed of 490 messages, with attachments out of total approximately 78,820 unsolicited e-mail messages received in an e-mail account

during June 2005 to March 2008. When the sample was taken, UEMAs constituted 19.2 % of all the 26,160 unsolicited e-mail messages. When the second sample was taken, they constituted 9.3 % of 56,750 unsolicited e-mail messages collected during the same period. Figure 1 gives a snapshot of the e-mail account, showing the quantity of messages.

The account has received 78,910 messages in total, with normal messages accounting for a small proportion and 78,820 unsolicited messages in folders “research “research1, “501 and “490. 501 UEMAs have been collected in folder “510. The 490 UEMAs have been collected in folder “490.” The first sample, which constituted the framework of this study, was analyzed in May 2006. The second sample, analyzed in March 2008, was primarily used to supplement the previous findings.



Name	Messages	Unread
Inbox	4	-
Draft	2	2
Sent (Optional)	0	-
Bulk (Enter - Optional)	0	-
Trash (Empty)	0	-
490 (Rename - Delete)	490	-
501 (Rename - Delete)	501	-
backup (Rename - Delete)	87	-
research (Rename - Delete)	6,330	60003
research2 (Rename - Delete)	17499	17447
Total	78910	77452

Figure 1 Snapshot of the e-mail account showing the quantity of messages (as of 30 March 2008)

Because of the private nature of this e-mail account, it is easy to judge which message is unsolicited. In analysing each message, it is necessary to establish a standard to categorise messages into different types according to their sender column, subject column, content and attachments. The standard to decide if the sender column is falsified is the name format. The name for an organization is also easy to judge by comparing the name in the sender column with that in the content.

The standard for deciding whether a subject column is falsified can be loosely defined. Because there is only one message labelled with an “AD:” sign, in strict sense all the subject columns of other messages are illegal. However, the emphasis of this paper is not to coincide with the legal standard. Rather, it is focused on analysis of phenomena of UEMAs. The message with “AD:” label and messages with words explaining the content or having apparent connection with the content are regarded as not falsified. Other messages with subject column irrelevant with the content, inducing recipient to open messages, is considered falsified. The content is categorised according to offers provided, and the nature of attachments. Analysis of both samples is presented in this paper. Sections about validity of sender column, subject column and content of UEMAs are primarily based on the first sample.

Limits of this study are that the resource account was not broadly put on the Internet, but was published on only one website specialised on traditional publishing service, to soliciting submission of academic articles. It is difficult to determine whether a random sample of all UEMAs sent in the stream of commerce would yield similar findings. It is also unknown that whether publishing- and printing-related UEMAs are due to the resource specialisation.

Findings

Types of File Formats of Attachments

In the first sample of 501 UEMAs, three attachments were missing. Other attachments were comprised of 14 kinds of document formats. More than one third of attachments were “zip” format compressed files, mostly viruses, which represented the most severe threats to the e-mail users’ computer security. Another one third of all attachments were comprised of two categories: approximately one fifth of the total attachments being “gif” format image files, and approximately one sixth being “htm” format documents. This one third was relatively virus free, but included annoying contents and hyper links. These three kinds of files constituted more than 70 percent of all attachments. Another frequent attachment format was Microsoft Word “doc”, which accounted for 8.4 percent of all attachments. Other 10 kinds of document formats were only responsible for approximately 20 percent of

attachments, including both viruses and virus free files.

Three of the second sample of 490 UEMAs lost their attachments. Other 487 messages had 496 attachments, which included more types of files than in the first sample. Text files and “gif” files accounted for nearly 70 percent of these attachments. The most obvious change happened in growing role of text files and declining role of “zip” files, from which I observed that the use of UEMAs had become more rational in advertising practical information other than in spreading viruses. This conclusion was also supported by the decrease of executable files or their disguised formats, such as “com,” “exe,” “pif,” “scr,” and so on. Another observable change was that “html” files decreased by nice percent in attachments. In both samples, “gif” files had a significant percentage, but in fact, they were usually small, benign and meaningless. In sum, types of attachments took on a diversified outlook, with files of familiar formats changing to rational advertisements, and malicious motives seeking unfamiliar formats, at a satisfactorily smaller overall scale.

Table 1 Types of File Formats of Attachments

(In second sample, some messages had multiple attachments)

	Types of File Form ats	Attachments in first sample		Attachments in second sample		Change of types Percentage
		Number	Percentage	Number	Percentage	
1	*.chm	11	2.2	4	0.8	-1.4
2	*.com	9	1.8	2	0.4	-1.4
3	*.doc	42	8.4	40	8.1	-0.3
4	*.exe	13	2.6	5	1.0	-1.6
5	*.gif	92	18.5	129	26	+7.5
6	*.htm	78	15.7	32	6.5	-9.2
7	*.jpg	13	2.6	12	2.4	-0.2
8	*.mid	3	0.6	0	0	-0.6
9	*.pif	19	3.8	1	0.2	-3.6
10	*.rar	11	2.2	16	3.2	+1.0
11	*.scr	12	2.4	4	0.8	-1.6
12	*.txt	8	1.6	216	43.5	+41.9
13	*.xls	3	0.6	2	0.4	-0.2
14	*.zip	184	36.9	12	2.4	-34.5
15	*.epf	0	0	1	0.2	+0.2
16	*.hqx	0	0	10	2.0	+2.0
17	*.ini	0	0	1	0.2	+0.2
18	*.pdf	0	0	1	0.2	+0.2
19	*.png	0	0	1	0.2	+0.2

20	*.rtf	0	0	2	0,4	+0,4
21	*.url	0	0	2	0,4	+0,4
22	*.uu	0	0	3	0,6	+0,6
	Mess ages with missi ng attach ment	3		3		
	Total attach ments	498	100	496	100	

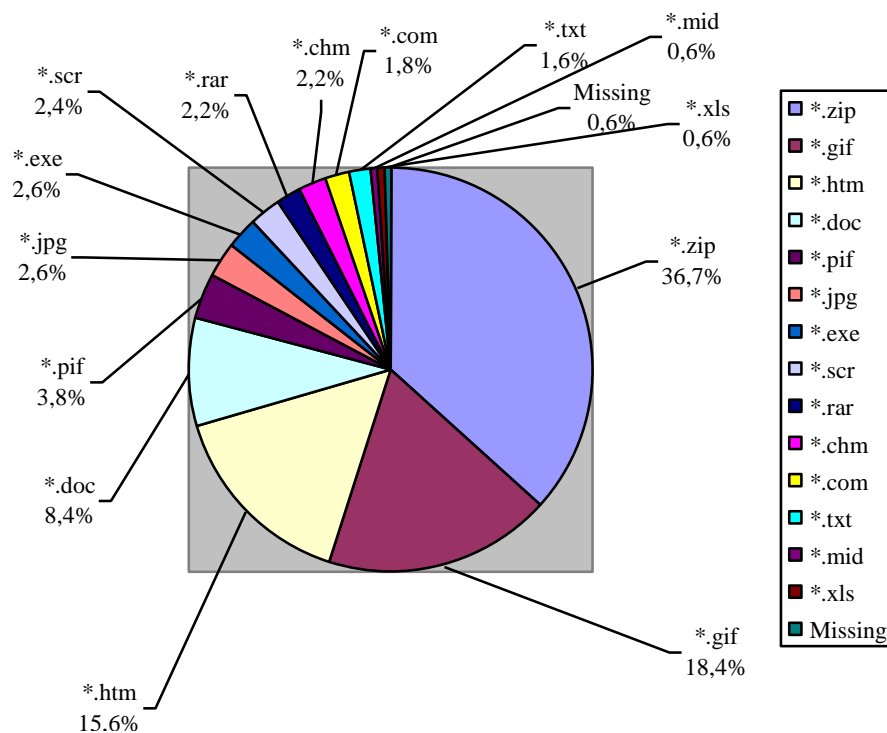


Figure 1 Types of File Formats of Attachments in 2006 Sample

Sizes of UEMAs

The average size of messages in the first sample was 47.44kb. Sizes of approximately 90 percent of UEMAs were smaller than 100kb. Sizes of only 2 percent of these messages were between 100-200kb. Sizes of nearly 9 percent of messages were bigger than 200kb. In fact, about 285 pieces of messages, which constituted more than half of UEMAs, were smaller than 30kb. Messages with the sizes of 1kb and 2kb alone accounted for more than 28 percent of the sample. They were mostly empty “zip” files with the possibility of being UEMAs of viruses

but disinfected by e-mail service providers. UEMAs spreading W32.netsky.C@mm (35k) and W32.Sober.X@mm (75) viruses accounted for 16 percent and 7 percent of all of messages respectively.

In the second sample, the average size was 76.53kb, about 61 percent bigger than that of the first sample. UEMAs smaller than 100kb constituted almost the same percentage as in the first sample. About three percent more messages were between 100 and 200kb, and fewer messages were bigger than 200kb. Nearly 70 percent of messages are smaller than 30kb. UEMAs smaller than 1kb and 2 kb apparently decreased in the second sample, accounting for only 6 percent.

Table 2 Sizes of UEMAs

	Massages in the first sample		Massages in the second sample	
Size	Numbers	Percentage	Numbers	Percentage
<100kb	446	89.0	424	86.5
100-200kb	11	2.2	26	5.3
>200kb	44	8.8	40	8.2
Total size				
23,765kb		37,500kb		
Average size	47.44kb		76.53kb	
Among which				
<30kb	285	56.7	377	76.9
1kb	13	2.6	3	0.6
2kb	108	21.6	25	5.1
35kb	80 (Virus: W32.Netsky.C @mm)	16		
75kb	35 (Virus: W32.Sober.X@ mm)	7		
40-45kb			3 (Viruses: W32.Blackmail. E@mm!enc, W32.Lovgate.R @mm)	0.6

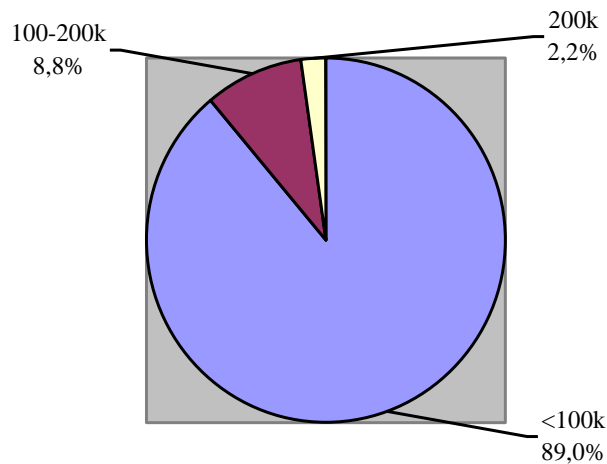


Figure 2 Sizes of UEMAs in 2006 Sample

In the first sample, UEMAs accounted for less than 2 percent of all unsolicited messages. The average size of unsolicited e-mails with attachments was 47.44kb, compared with the average size of 7.32kb of other unsolicited e-mails, a 6.5 fold bigger. In fact, UEMAs contributed to more than 10 percent of the average message size of all unsolicited messages, enlarging the average size from 7.32kb of unsolicited messages without attachments to 8.09kb of all of the unsolicited messages. When calculating message sizes in the second sample, I found that the average size had significantly enlarged, since there were 8 messages with the size ranging from 1mb to 4mb. These messages alone contributed to an average of 31.8kb for all the messages in the sample. Interestingly, large-sized messages were usually sent by political-oriented groups with the intent to spread political speeches.

Table 3 Compare of Average Sizes of Messages in 2006 Sample

	Total size	Numbers	Average size
Average size of unsolicited e-mails with attachments	23,765k	501	47.44k
Average size of other unsolicited e-mails	186,254k	25,459	7.32k
Average size of all unsolicited e-mails	210,019k	25,960	8.09k

Types of Sender Columns in UEMAs

Among the 501 pieces of UEMAs, one is with a blank sender column. Senders of UEMAs tend to hide their names but show e-mail addresses, valid or false. Approximately 60 percent of all messages show e-mail addresses instead of senders' name, which should be considered substandard. Others conceal their names with their surnames and titles, meaningless letters and numbers, describing their offers (products, services and activities), filled with words inducing users to open messages, or simply exploiting recipients' names and e-mail address. In total, approximately 83 percent of messages bear substandard sender columns. Only around one sixth of messages bear standard personal names or company names.

Table 4 Types of Sender Columns in UEMAs

Types of sender columns	First sample		Second sample		Changes in percentage
	Number	Percentage	Number	Percentage	
Blank	1	0.2	1	0.2	0
Company name	45	0.9	27	5.5	+4.6
Describing products, services and activities	16	3.2	136	27.8	+24.6
Inducing users to open messages	7	1.4	14	2.9	+1.5
Meaningless letters and numbers	23	4.6	105	21.4	+16.8
Recipients' name and address	4	0.8	1	0.2	-0.6
Showing e-mail address	297	59.3	72	14.7	-44.6
Standard personal name	79	15.8	86	17.6	+1.8
Surname plus title	29	5.8	48	9.8	+4.0
Total	501	100	490	100	

In the second sample, messages with sender columns describing products, services and activities increased by a quarter. Messages with sender columns comprised of meaningless letters and numbers increased by nearly 17 percent. Showing e-mail addresses in sender columns decreased nearly 47 percent.

Valid Sender Columns in UEMAs

With the first sample, I analyzed in more details validity of sender column, subject column, and content. The following sections are primarily based on the first sample. Senders of UEMAs that were currently empty, or with the content of offering banking or financial services, sales of falsified certificate, human resources recruitment, publishing and printing, sales of health products and clothes, soliciting friends, and with the purpose of merely spreading computer viruses were reluctant to provide sender names in standard formats. Senders of UEMAs who offered telecommunications services were also quite reluctant to do so. Approximately one in every three senders of UEMAs offering information on companies and websites, sales of books, VCD and DVD provided valid name format in sender column. Half of quick money opportunities providers typed right names in their messages' sender column. Senders of UEMAs that offered tax evasion assistance seemed more active in providing standard format of names in the sender column. More than half of them did so. Sixty percent of senders who offered information on training and education opportunities typed names in standard format in the sender column. The providers of computer hardware and software appeared the most reliable senders of UEMAs, of whom more than 70 percent furnished standard sender column. One quarter out of senders who offered other services gave valid form of names.

Table 5 Valid Sender Columns in UEMAs in 2006 Sample

Type	Number of validity in sender column	Percentage
Banking, financial	0	0
Empty attachments	0	0
Falsified certificate	0	0
Human resources recruitment	0	0
Introduction of company, website	6	33.3
Political propaganda	14	56
Publishing, printing, card manufacture, etc.	0	0
Quick money	1	50
Sales of books, VCD, DVD	4	36.4
Sales of health products, clothes	0	0
Software, computer products	33	70.2
Soliciting friends	0	0
Tax evasion	37	53.6
Telecommunications services	1	3.4

Training and education	6	60
Virus	0	0
Other services	1	25

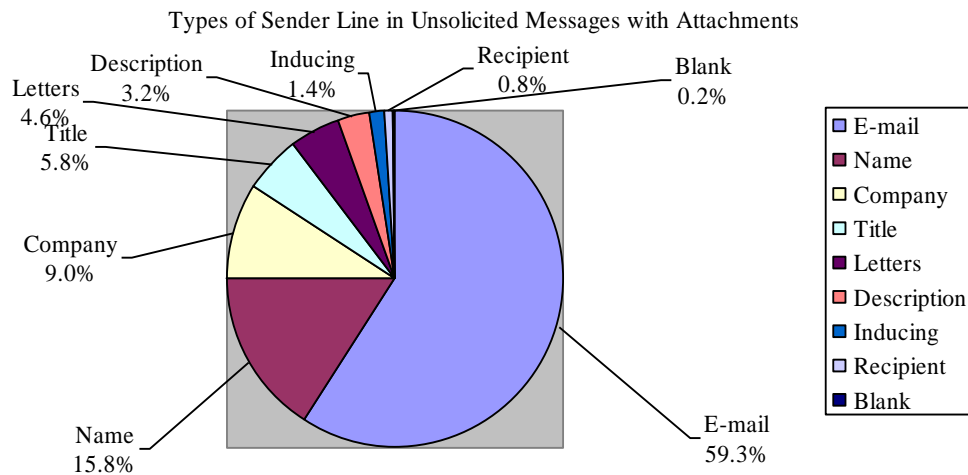


Figure 3 Types of Sender Columns in UEMAs in 2006 Sample

Types of Subject Columns in UEMAs

To label the subject column with “AD:”, “ADV:”, or any other kinds of regulatory means, was an invention that had never been respected by senders of unsolicited messages. According to calculation of the first sample, only less than two in one thousand messages had this kind of label. Two in one hundred of UEMAs left the subject column blank. More than 17 percent of messages used ambiguous wording to confuse recipients, or other particular terms attempting to draw recipients’ attention and attracting them to open messages, furnished the subject column with languages describing the content, giving greetings, appearing related to users’ e-mail service, bearing “To:”, “Re:” and “Fw:” labels, or pretending users’ friends and contacts, etc. The hacking tactics of so-called social engineering was to a great extent used in these messages. In the second sample, messages with subject columns describing message content increased by nearly 50 percent, while messages with subject columns pretending to be recipient's contact decreased by more than 40 percent.

Table 6 Types of Subject Columns in UEMAs

	First		Second		Change in
--	-------	--	--------	--	-----------

	sample		sample		percentage
Type	Number	Percentage	Number	Percentage	
“AD:” label	1	0.2	0	0	-0.2
Attractive wording	87	17.4	26	5.3	-12.1
Blank	10	2	8	1.6	-0.4
Describing message content	117	23.4	354	72.2	+48.8
“To:”, “Re:” and “Fw:” label	21	4.2	25	5.12	+0.92
Users’ friends and contacts	265	52.9	44	8.98	-43.92
Other	0	0	33	6.7	+6.7
Total	501	100	490	100	

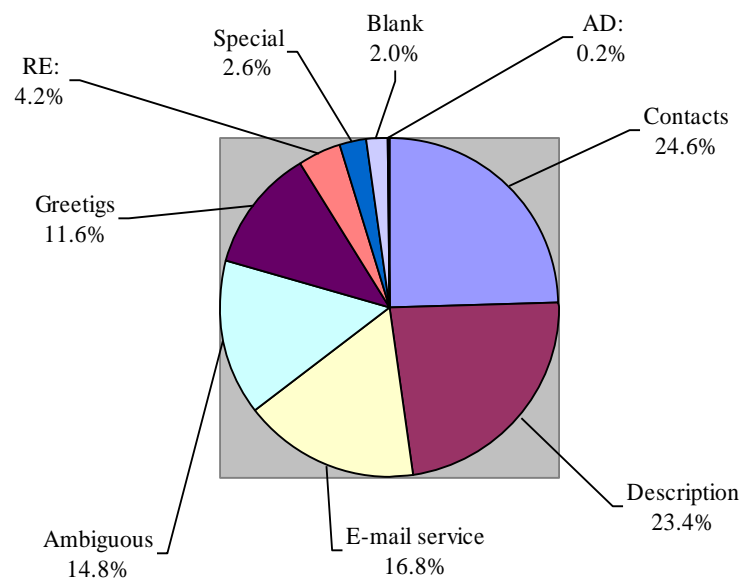


Figure 4 Types of Subject Column in UEMAs in 2006 sample

Valid Subject Columns in UEMAs

Almost all of senders of UEMAs offering information on human resources recruitment, companies or websites, publishing, printing, card manufacture, etc., sales of health products, clothes, and tax evasion ensured the valid subject column. A high percentage of providers of telecommunications services, sellers of books, VCDs, and DVDs, training information providers, quick money information providers, and providers of other services provided valid subject columns in their messages. All senders of other messages were reluctant in typing

useful subjects for their potential recipients.

Table 7 Valid Subject Columns in UEMAs in 2006 Sample

Type	Number	Percentage
Bank, financial	0	0
Empty attachments	11	10.7
Falsified certificate	0	0
Human resources recruitment	2	100
Introduction of company, website	18	100
Political	1	4
Publishing, printing, card manufacture, etc.	19	100
Quick money	1	50
Sales of books, VCD, DVD	8	72.7
Sales of health products, clothes	4	100
Software, computer products	1	2
Soliciting friends	0	0
Tax evasion	66	95.6
Telecommunications services	24	82.8
Training	7	70
Virus	0	0
Other services	3	75

Types of Content of UEMAs

More than 28 percent of messages were designed to spread viruses. More than one in five messages attached empty attachments. Messages offering both tax evasion services and software and computer products constituted around 10 percent of all messages. Any of contents of other messages constituted a percentage far below 10 percent, with messages offering telecommunications services and political propaganda constituting around 5 percent separately. Interestingly, there was rarely any message with attachment involving adult contents, investment chances, sales of pirated software, and some other common offers in messages without attachments.

Table 8 Types of Content of UEMAs in 2006 Sample

Type	Number	Percentage
------	--------	------------

Bank, financial	3	0.6
Empty attachments	103	20.6
Falsified certificate	10	2
Human resources recruitment	2	0.4
Introduction of company, website	18	3.6
Political	25	5
Publishing, printing, card manufacture, etc.	19	3.8
Quick money	2	0.4
Sales of books, VCD, DVD	11	2
Sales of health products, clothes	4	0.8
Software, computer products	47	9.4
Soliciting friends	4	0.8
Tax evasion	69	13.8
Telecommunications services	29	5.8
Training	10	2
Virus	141	28.1
Other services	4	0.8

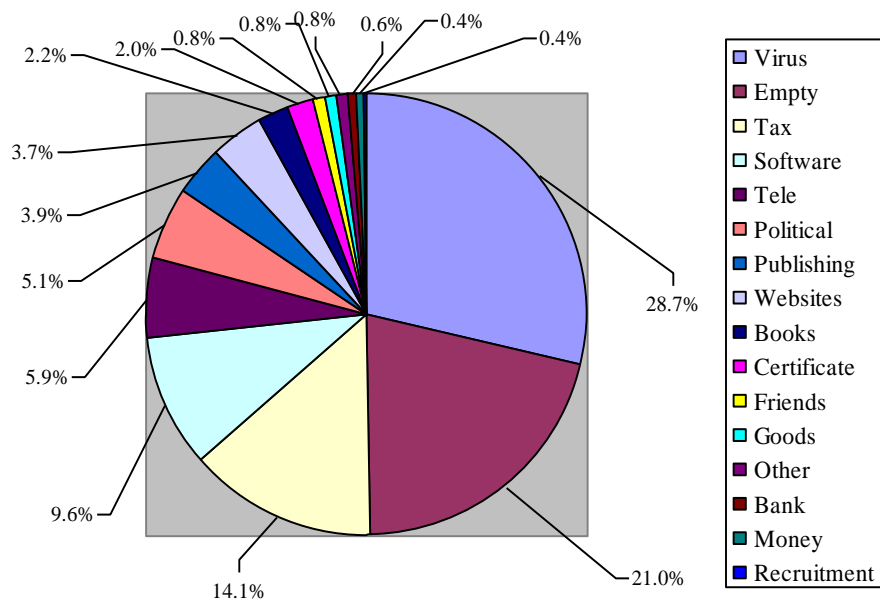


Figure 5 Types of Content of UEMAs in 2006 Sample

Valid Content in UEMAs

Interestingly, most messages had high percentage of valid contents, with exception of messages with empty attachments and UEMAs that spread viruses.

Table 9 Valid Content in UEMAs in 2006 Sample

Type	Number of validity in content	Percentage
Bank, financial	0	0
Empty attachments	0	0
Falsified certificate	9	90
Human resources recruitment	2	100
Introduction of company, website	13	72
Political	25	100
Publishing, printing, card manufacture, etc.	19	100
Quick money	1	50
Sales of books, VCD, DVD	11	100
Sales of health products, clothes	4	100
Soliciting friends	4	100
Software, computer products	47	100
Telecommunications services	24	82.8
Training	8	80
Virus	0	0
Other services	1	25

Types of Contact Methods Provided in UEMAs

Because many UEMAs were spreading viruses, they generally provided no contact information, with a few exceptions. Other messages included one or more kinds of contact methods in the message texts. Of total 501 messages in the first sample, more than one-third of messages provided hyperlinks directed to websites, while less than one-third provided fixed telephone numbers. Both e-mail addresses and mobile phone numbers were preferred by

more than 22 percent of senders. Fax numbers and physical addresses were included in about 16 and 10 percent of messages separately. QQ (a chat system) were provided in 8 percent of messages. MSN was the least used contact method in the 501 messages.

Unsubscribe is nothing more than a decoration in UEMAs. Unsubscribe method was only provided in 2.2 percent of messages in the first sample.

Table 10 Types of Contact Methods Provided in UEMAs in 2006 Sample

Contact Methods	Number	Percentage (/501)
Address	50	10
E-mail	113	22.6
Fax	79	15.8
MSN	12	2.4
Mobile phone	112	22.4
QQ	40	8
Telephone	145	28.9
Unsubscribe method	11	2.2
Website	182	36.3

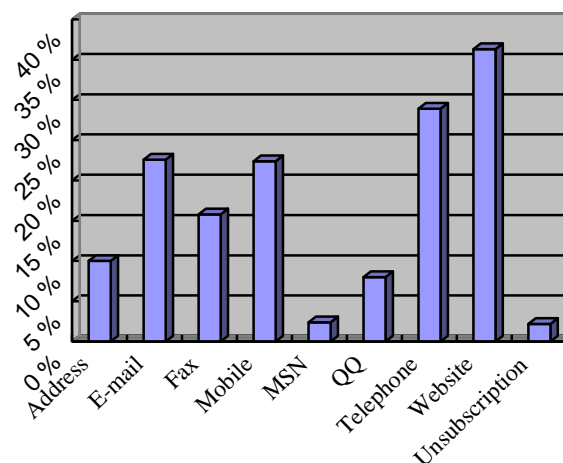


Figure 6 Types of Contact Methods Provided in UEMAs in 2006 Sample

Falsity of Sender, Subject and Content

In the first sample, only one in five of UEMAs used valid format in sender column, one in three used valid subject column, and more than half provided valid content. However, only 58 messages had both valid format of sender and subject column, and 293 messages with both

false format of sender and subject column. Other 42 messages had valid format of sender column but false format of subject column, while 108 messages had false format of sender column but valid format of subject column. The overall falsity of subject or sender column constituted 88.4 percent.

Table 11 Validity and Falsity of Sender and Subject Columns and Content Separately

		Number	Percentage
Sender	Valid	100	20
	False	401	80
Subject	Valid	166	33.1
	False	335	66.9
Content	Valid	260	51.9
	False	241	48.1

The validity and falsity of content took almost 50 percent separately.

Table 12 Validity and Falsity of Either Subject or Sender Columns

		Sender			
		Valid	False	Total numbers	Percentage
Subject	Valid	58	108	166	33.1
	False	42	293	335	66.9
	Total numbers	100	401	501	100
	Percentage	20	80	100	
	Percentage of validity of both columns	11.6			

Discussion

In activities of sending unsolicited e-mail messages, one of the most important aspects is to induce recipients to open and read messages and their attachments. Senders of UEMAs took particular considerations in disguising their real identity and real purpose. It can be said that most of them demonstrated a high degree of skill in motivating recipients to open messages and their attachments. However, opening messages and attachments was usually the first step towards for senders to victimize other users. Generally, they used ambiguous wording in sender and subject columns but valid content (except messages spreading viruses)

so as to ensure messages and attachments be open and advertisements be read.

Analysis of the two samples roved that, except messages deliberately spreading viruses, UEMAs were usually less harmful than it was found in findings of previous studies that did not distinguish UEMAs from those without. It implied that senders of unsolicited e-mail messages tended to transform their product or service information.

At the same time, this study revealed that UEMAs could have broader negative impact on criminal phenomena, not only victimization, but also conspiracy. E-mail communications could be taken as an offensive means by which recipients were victimized, or a conspiracy tool with which recipients were seduced to commit crimes. UEMAs that spread viruses could directly result in victimization of recipients, their computers damaged or manipulated, or their secret or privacy stolen or disclosed. Some UEMAs could move recipients to invest their money to projects that could never generate any reward. Some UEMAs could victimize recipients in crimes that recipients being physically or sexually attacked. Or, these messages could induce recipients to collaborate on some criminal plots.

In cyber environment, the most frequent victimization model began from exposing of victims to potential threats, which we can label as exposing-victimization model. Under this model, the victim of UEMAs exposed their addresses merely on web pages, in bulletin board systems (BBSes), in chat systems, or simply in transmission through the Internet. Exposure on the Internet does not necessarily mean show-off. Rather, it should have been a usual kind of digital presence. Nevertheless, the exposing-victimization model at least implies that senders of UEMAs could easily harvest e-mail addresses in the same way as normal Internet users do

In other cases, senders of UEMAs had a seeking process, and followed the seeking-victimization model. Due to the large quantity of web pages and other Internet-related contents, direct artificial collection of multiple e-mail addresses became inefficient. By making use of specialized software to harvest e-mail addresses from the Internet, senders could collect millions of addresses automatically in a short time. By doing so, they created the seeking-victimization model in sending UEMAs. Besides harvesting, they also exploited dictionary attack and/or automatic alphabetical permutation and combination to enumerate possible usernames in e-mail accounts. All of these methods could be used in seeking process. For senders, an e-mail account with a random word might not represent a specified person;

but for the recipient, he/she would readily be the potential victim of this UEMA.

Victimization of recipients of UEMAs could happen without recipients accessing their e-mail accounts. Their victimization resulted from their e-mail accounts being spammed, whether they accessed their accounts or not. Under current legal framework, receiving of unsolicited messages is sufficient to be regarded as victimization caused by acts that imposed punishment by law.

However, victimization of UEMAs does not end at the initial victimization. The above-mentioned models could be called as the first level effects of UEMAs. The second level effects could take place on basis of the initial victimization. There could also be two sub-models: victimization-victimization sub-model and victimization-conspiracy sub-model.

The victimization-victimization model happened when messages spread viruses, fraudulent sales of goods, or falsified financing and banking services. The first level victimization was for recipients to be spammed, while the second level victimization was for recipients to be attacked or swindled upon opening the malicious programs or following the fraudulent scams.

The second level victimization would not always be accomplished so straightforwardly. Usually involved was a victimization-exposing-seeking-victimization process. The most typical scam of this kind was Nigerian fraud (or 419 fraud), in which recipients of unsolicited messages were firstly victimized by receiving this kind of messages (being spammed). If they made a positive reaction to messages, they were further exposing their vulnerabilities to senders. Upon receiving reply from recipients, senders would further seek vulnerabilities of recipients and at last obtain their property. The process of seeking and exposing might be a long interaction between senders and recipients, who exchanged messages until the final transaction. If senders succeeded in obtaining recipients' property, the last victimization would take place and the scam would come to an end.

The victimization-conspiracy model happened when messages included assistance for tax evasion services, sales of pirated software, sales of falsified documents, and so on. Recipients of such offer were firstly victimized by unsolicited messages; and if they participated in illegal operations, they would become conspirators of senders.

Because recipients of unsolicited messages inducing conspiracy in an illegal operation

would expect to have the luck to benefit from collaborating with senders who pretended to have the potential to give charity, senders were more likely to send this kind of messages. In fact, in Nigerian fraud, senders were usually personating politicians who want to transfer property (money, diamond, and so on) to bank accounts of recipients by claiming to give a significant reward. As a result, recipients, who wished they could have been “conspirators” in a great operation of money laundering, would finally be victimized in scams that they lost advance fees.

Conclusion

Phenomena of UEMAs further proved the low controllability or uncontrollability of cyber environment. Exposed e-mail addresses are vulnerable to unsolicited messages. Unexposed e-mail addresses can be as vulnerable as exposed ones, because address harvesting software can collect potential addresses from transmission route of the Internet. As far as our study is concerned, this process makes extra sense. For senders, both ways could be seen as a process of seeking vulnerabilities. For recipients, both ways could also be seen as a process of exposing vulnerabilities. However, seeking and exposing process has become more abundant and colourful in cyber environment than pre-Internet times.

Mere browse of web pages is the easiest method to obtain e-mail accounts, but it is less efficient because e-mail addresses are usually scattered in many different pages. This process would be time-consuming if thousands or millions of addresses are to be collected. Senders can also purchase millions of addresses of users with different interests from specific vendors, who have the best method and specialized personnel to harvest addresses from all over the cyberspace, and usually establish their own databases of addresses. With an inexpensive price, buyers can conveniently get a large quantity of addresses. In addition, address harvesting becomes automated and prevalent with the help of specified software. Many people who are interested in doing spamming business can easily master uncomplicated skills and collect millions of addresses with such software, which can be downloaded from the Internet free of charge or with a small payment.

Exposing e-mail addresses on the Internet is literally unavoidable, because the exposure

is in so broad a sense that all the normal use of e-mail services could be seen as an exposing process, including sending and receiving messages; publishing on web pages, chat rooms, and BBSes; providing as register information in online services; or exposing nothing but coincidence with a dictionary vocabulary; and so on. In fact, exposing a single e-mail account will not be so risky if there is not such a thing as address harvesting technique, because it is an inefficient way to collect single e-mail account one by one from the Internet. However, we cannot simply ignore such a method because e-mail account vendors could collect and transact addresses in a dynamic process, and collect addresses through a variety of ways to establish their databases. Address harvesting software and dictionary attack undoubtedly intensify the victimization of e-mail account holders.

In general, exposed e-mail accounts might face double risks of being victimized: being collected in process of formally browsing web pages and use of other Internet services; and being harvested during the process of digital transmission or merely guessed by senders through randomly combining letters and numbers. Compared with daily used e-mail accounts without exposing on web pages or other Internet services, published accounts are more likely to be spammed. Therefore, it seems more likely that vendors or senders harvest addresses with automated technique. As a result, double risks of exposed e-mail accounts are in fact unbalanced: the risk of being victimized by collectors and harvesters are far serious than that by guessers.

UEMAs provide e-mail users many different choices, either conspiring in criminal acts, or victimized by viruses or in scams. Messages analyzed in this study generally gave recipients two alternatives: conspiring in tax evasion, or damaged by viruses.

In the case of conspiracy in tax evasion, senders used to provide valid contact methods so as to induce recipients to participate in illegitimate operations. The offer seemingly aims to establish a relationship between tax evasion service provider and their potential clients. However, the effect was that they form conspiracy in tax evasion activity. Recipients had to react actively before they become conspirators of tax evasion activities. The process might involve repeated e-mail exchanges upon initial unsolicited messages. Under these circumstances, unsolicited messages might be transformed into literally valuable (but morally wrong and legally prohibited) information for recipients. Thus recipients might tend to accept

such messages and offers in them. Such messages become the communication means for criminals, posing great threats for social control over illegal activities.

In the case of viruses attack, senders exploited social engineering to induce recipients to open messages and their attachments, by blurring sender, subject columns and falsifying message content and file names of attachments. These messages did not require any reply from recipients before they caused damages. They were also dangerous for recipients in the sense that they were harming recipients' hardware and software, wasting time and labour.

References

Boldt, M., Carlsson, B., & Jacobsson, A. (2004). Exploring Spyware Effects. Retrieved April 1, 2008, from <http://www.tml.tkk.fi/Nordsec2004/Presentations/boldt.pdf>

Cobb, S. (2003). The Economics of Spam. Retrieved April 1, 2008, from http://www.spamhelp.org/articles/economics_of_spam.pdf

Fallows, Deborah (2003, October). Spam: How It Is Hurting E-mail and Degrading Life on the Internet. Retrieved April 1, 2008, from http://www.pewinternet.org/pdfs/PIP_spam_Report.pdf

Federal Trade Commission (1998, July). Federal Trade Commission Names Its Dirty Dozens: 12 Scams Most Likely to Arrive via Bulk E-mail, *Federal Trade Commission Consumer Alert*. Retrieved April 1, 2008, from <http://library.findlaw.com/1998/Jul/1/128450.html>

Federal Trade Commission (2002a, April), *Remove Me Surf*, Author. Retrieved April 1, 2008, from <http://www.ftc.gov/bcp/online/edcams/spam/pubs/removeme.pdf>

Federal Trade Commission (2002b, November), *E-mail Address Harvesting: How Spammers Reap What You Sow*, Author. Retrieved April 1, 2008, from <http://library.findlaw.com/2003/Aug/8/132973.pdf>

Federal Trade Commission (2003a, April), *False Claims in Spam: A Report by the Federal Trade Commission's Division of Marketing Practices*, Author. Retrieved April 1, 2008, from <http://www.ftc.gov/reports/spam/030429spamreport.pdf>

Federal Trade Commission. (2003b, June 15). *National Do-Not-E-mail Report to Congress*, Author. Retrieved April 1, 2008, from <http://www.ftc.gov/reports/dneregistry/report.pdf>

Gauthronet, S., & Drouard, E. (2001). *Unsolicited Commercial Communications and Data Protection*. Brussels: Commission of the European Communities, Internal Market Directorate General. Retrieved April 1, 2008, from http://ec.europa.eu/justice_home/fsj/privacy/docs/studies/spamsum_en.pdf

Goodman, Danny (2004), *Spam Wars: Our Last Best Chance to Defeat Spammers, Scammers & Hackers*, New York, New York: SelectBooks.

Goodman, J. T., and Rounthwaite, R. (2004). Stopping Outgoing Spam. In: *Proceedings of the 5th ACM Conference on Electronic Commerce*, 17-24 May, ACM Press, pp. 30-39. Retrieved April 1, 2008, from <http://research.microsoft.com/~joshuago/outgoingspam-final-submit.pdf>

IDC (2005, February 24). Worldwide Revenue for Antispam Solutions To Reach Over \$1.7 Billion in 2008, IDC Reveals. *IDC - Press Release*. Retrieved April 1, 2008, from http://findarticles.com/p/articles/mi_m0EIN/is_2005_Feb_24/ai_n10300118

Karnell, J. (2002). Raising the Stakes in Permission Marketing. Retrieved April 1, 2008, from <http://www.onetooneinteractive.com/resource/whitepapers/0003.html>

Khong, W. K (2001, October). The Law and Economics of Junk E-mails (Spam). Retrieved April 1, 2008, from <http://www.emle.org/Thesis/Khong.pdf>

Khong, W. K. (2004). An Economic Analysis of Spam Law. *Erasmus Law and Economics Review*, 1 (February), 23–45. Retrieved April 1, 2008, from <http://www.eler.org/include/getdoc.php?id=8&article=2&mode=pdf&OJSSID=6170ccc598edb033fc0ccf2477a86ee9>

Lambert, Anselm (2003, September). *Analysis of Spam*. Master of Science in Computer Science Dissertation, Dublin: University of Dublin.

Li, Xingan. (2006). E-marketing, Unsolicited Commercial E-mail, and Legal Solutions, *Webology*, 3(1), Article 23. Retrieved April 1, 2008, from <http://www.webology.ir/2006/v3n1/a23.html>

Li, Xingan (2007). The Phenomenon of Unsolicited E-mails with Attachments. *SIMILE: Studies In Media & Information Literacy Education*, 7 (2), 1–11.

McWilliams, Brian (2005). *Spam Kings*, Sebastopol: O'Reilly Media.

Nucleus (2003). *Spam: The Silent ROI Killer*, Research Note D59. Retrieved April 1, 2008, from <http://www.spamhelp.org/articles/d59.pdf>

Nucleus (2004). *Spam: The Serial ROI Killer*, Research Note E50. Retrieved April 1, 2008, from http://tim.blog.kosmo.com/article_files/NucleusResearchCostOfSpam.pdf

Organization of Economic Cooperation and Development (2003). *Organization of Economic Cooperation and Development Guidelines for Protecting Consumers from Fraudulent and Deceptive Commercial Practices Across Borders*, Author. Retrieved April 1, 2008, from http://www.oecd.org/document/56/0,2340,en_2649_34267_2515000_1_1_1_1,00.html

Organization of Economic Cooperation and Development (2004). *Second Organization of Economic Cooperation and Development Workshop on Spam: Report of the Workshop*, Author. Retrieved April 1, 2008, from <http://www.oecd.org/dataoecd/55/32/31450810.pdf>

PC World (2003, August 29). Sobig May Be Working for Spammers. Retrieved April 1, 2008 from <http://www.pcworld.com/news/article/0,aid,112261,00.asp>

Radical Group (2005). *The Radical Group, Inc. Release Q1 2005 Market Numbers Update*, Author. Retrieved April 1, 2008, from http://www.radicati.com/uploaded_files/news/Q1-2005_PressRelease.pdf

Sadowsky, G., Dempsey, J. X., Greenberg, A., Mack, B. J., and Schwartz, A. (2003). *Information Technology Security Handbook*. The International Bank for Reconstruction and Development.

Simon, H. A. (1982). *Designing Organizations for an Information-Rich World: Models of Bounded Rationality*. Massachusetts: Massachusetts Institute of Technology Press.

Sorkin, D. E. (2001). Technical and Legal Approached to Unsolicited Electronic E-mail, *University of San Francisco Law Review*, Vol. 35, pp. 325-384. Retrieved April 1, 2008, from <http://www.sorkin.org/articles/usf.pdf>

Spammer-X (2004), *Inside the SPAM Cartel*, Rockland, Massachusetts: Synergies Publishing.

Taylor, Humphrey (2003, 10 December). Spam Keeps on Growing. Retrieved April 1, 2008, from http://www.harrisinteractive.com/harris_poll/index.asp?PID=424

Trans Atlantic Consumer Dialogue (TACD) (2003). *Consumer Attitudes Regarding Unsolicited Commercial E-mail (Spam)*, Author. Retrieved April 1, 2008, from http://www.tacd.org/db_files/files/files-296-filetag.doc

World Summit of Information Society (2003, December). *Declaration of Principles-Building the Information Society: A Global Challenge in the New Millennium*, Author. Retrieved April 1, 2008, from http://www.itu.int/dms_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0004!!PDF-E.pdf

Federal Trade Commission (2005). Email Address Harvesting and the Effectiveness of Anti-Spam Filters: A Report by the Federal Trade Commission's

Division of Marketing Practices, November 2005, 10 pp. Retrieved April 1, 2008, from <http://www.ftc.gov/opa/2005/11/spamharvest.pdf>